Pergamon

Neural Networks, Vol. 8, No. 7/8, pp. 1215–1226, 1995

0893-6080/95 $9.50 + .00

0893-6080(95)00069-0

1995 SPECIAL ISSUE

Biologically Motivated Cross-modality Sensory Fusion System for Automatic Target Recognition

TERRY HUNTSBERGER

University of South Carolina

(*Received* 28 *October* 1994; *revised and accepted* 11 *April* 1995)

Abstract—*The need for robust target/background segmentation has led to the use of multiple band sensing systems. These sensors usually include some combination of visual, radar or laser range, and thermal infrared modalities. Despite over a decade of research, there are still a number of problem areas with existing automatic target recognition systems. Foremost among these are the high false-alarm rates frequently encountered due to nonrepeatability of the target signatures and possible obscuration of the targets from camouflage, environmental and sensor variations (Roth, 1989, IEEE Transactions on Systems, Man and Cybernetics, 19, 1210–1217; Roth, 1990, IEEE Transactions on Neural Networks, 1, 28–43). This paper presents a biologically motivated neural network system based on the rattlesnake that integrates multichannel sensory inputs for ATD/R. The system demonstrates a probability of detection greater than 90% with false-alarm rate less than* 10^{-5} *false-alarms/km*^{2} *for very small fixed targets using two-channel infrared input. In addition, temporal properties of the thermal neurons in the rattlesnake are demonstrated to be of possible use for segmentation of mobile targets from background clutter. Also presented are the results of some experimental studies on real-world multichannel infrared images sampled throughout a day.*

Keywords—ATR, Neural network, Rattlesnake, Sensor fusion, Fuzzy sets.

1. INTRODUCTION

Traditional statistical approaches to the ATR problem are often hampered by changes in environmental conditions. A high probability of detection (P_D) is sometimes accompanied by high false-alarm rates (FAR). Neural networks offer an alternative choice for ATR applications, due to their inherently adaptive and parallel properties (Roth, 1989, 1990). Some recent work along these lines with radar, infrared and sonar signals have demonstrated a P_D that is usually greater than 90% and a FAR that is negligible (Gorman & Sejnowski, 1988; Farhat & Bai, 1989; Landowski & Fong, 1990; Bai & Farhat, 1992; Hemminger & Pao, 1994).

Biological systems are optimized for ATR-like applications in the predator–prey relationship and lend themselves well to hardware implementations (Hartline et al., 1978; Schiff et al., 1994). In addition, biological systems offer real-time solutions to target localization and tracking problems (Grossberg et al., 1993). There are numerous biological neural network systems that integrate sensory information for an enhanced view of the environment. Among these are the barn owl (Spence et al., 1989; Spence & Pearson, 1990), the vestibulo-ocular reflex system (Paulin et al., 1989), the rattlesnake (Newman & Hartline, 1981) and the echolocating bats (Simmons, 1990). The optic tectum in the rattlesnake directly integrates bimodal inputs (visual and thermal infrared) in a set of six specialized neurons (Newman & Hartline, 1981). Our previous studies have indicated that these neuronal types can be used as generalized fusion filters for any pair of disparate sensory inputs (Huntsberger, 1990, 1992a, b). This includes both target signature determination and target discrimination tasks.

Feature selection for robust ATR is directly tied to target discrimination from the sensor input and

Acknowledgements: The author would like to thank Ernesto Cespedes of the Waterways Experiment Station for supplying the images used in the experimental study, and the reviewers, whose comments helped to clarify some issues in the paper. I would also like to thank Dr. Patrick Fay at the University of South Carolina for providing the Intel Paragon time for the experimental studies. This work was supported in part under ARO grant DAAL03-91-C-0034, AFOSR grant F49620-93-1-0083, and ONR grant N00014-94-1-1163. The views, opinions, and/or findings contained in this paper are those of the author and should not be construed as an official Department of Defense position, policy, or decision,unless so designated by other documentation.

Requests for reprints should be sent to T. Huntsberger, Department of Computer Science, University of South Carolina, Columbia, SC 29208, USA; E-mail: terry@cs.sc.edu.

any subsequent processing. Uncertainties and ambiguities in single sensor responses can be resolved using multiple sensors (Perlovsky, 1987; Ajjimarangsee & Huntsberger, 1988; Nandhakumar & Aggarwal, 1988; Huntsberger, 1990; Chu & Aggarwal, 1992). Fuzzy set theory approaches, when combined with sensor fusion, further reduce the effects of such sensor limitations as saturation (Chatterjee & Huntsberger, 1988; Perlovsky & McManus, 1991; Huntsberger, 1992a, b; Krishnapuram & Lee, 1992). The primary advance made with the fuzzy set theory approaches lies in the rich level of representation built into the framework (Huntsberger et al., 1986). Variations in sensor response over a surface and edge strength degradation due to motion are directly included in the membership values to the image subsets. In addition, extensions for texture and motion analysis are straightforward (Huntsberger & Jayaramamurthy, 1987a, 1988; Huntsberger et al., 1987).

A generalized sensor input processing model is the key to the minimization of environmental change impact on ATR algorithm performance. Relative insensitivity to diurnal variations is built into our algorithm through the exploitation of cross-modality sensor interactions. These types of cross-modality fusion characteristics are found in specialized neurons in the optic tectum of the rattlesnake (Newman & Hartline, 1981). The cross-modality neuronal responses are simple enough to be simulated using lookup tables for efficiency (Ajjimarangsee & Huntsberger, 1988).

Our previous experimental studies compared the Dempster–Shafer (Shafer, 1976) and modified Dempster–Shafer (Ishizuka, 1981; Huntsberger & Jayaramamurthy, 1987b) statistical evidence combination methods to a neural network model based on the rattlesnake optic tectum for integration of sensor inputs (Chatterjee & Huntsberger, 1988; Huntsberger, 1992a). The results indicated that the biologically motivated neural network system performed as well as, or better than, the statistical methods. The neural network system was an order of magnitude faster since there was none of the overhead associated with the global normalization step of the Dempster–Shafer algorithm. These properties make the neural network system an ideal candidate for real-time ATR applications.

This paper presents an ATR system, shown in Figure 1, that is based on a three stage neural network model which contains fuzzy self-organizing feature maps combined with filters derived from the cross-modality fusion neurons in the optic tectum of the rattlesnake. These cross-modality fusion neurons give the system the ability to detect small fixed targets with a high probability, despite very large diurnal variations in the multi-channel infrared sensor inputs.

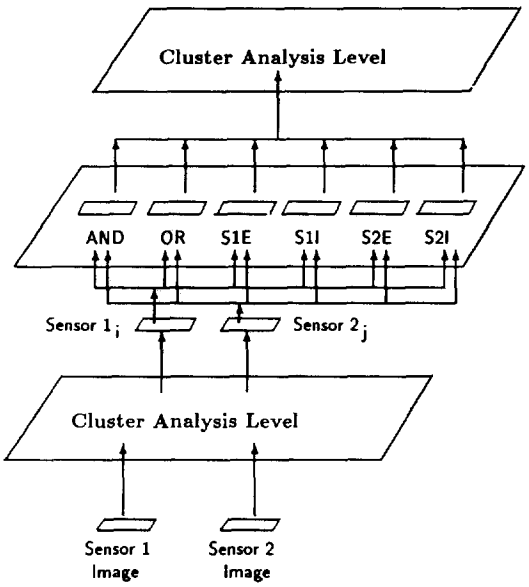

FIGURE 1. Multi-level neural network model for sensor integration.

Rattlesnakes exploit this capability for hunting in the late afternoon, when the thermal infrared pits are almost saturated from the landscape that has been heating all day.

The next section discusses some previous experimental biological background work and the overall neural network system design. Some properties of the fuzzy self-organizing feature map algorithm for unsupervised target signature determination are also included. This is followed by a discussion of extensions to the system for mobile target discrimination. Finally, the results of some experimental studies on both fixed and mobile targets using real-world multichannel infrared data are presented.

2. NEURAL NETWORK MODEL

Snakes in the subfamily Crotalinae all posses pit organs that are sensitive to heat through a dense network of nerve fibers. These pit vipers also have specialized routing and processing centers, which control information transfer to the optic tectum portion of the midbrain. Although the optic nerve fibers feed almost directly into the optic tectum, only the trigeminal nerve from the pit organ passes through these centers, which are called the LTTD and the reticularis caloris (RC). This process results in a type of spatial registration of the visual and thermal infrared inputs within the optic tectum.

An extensive study of single thermal neuron properties in pit vipers was done by Bullock and Diecke (1956). They observed that a low intensity input caused a tonic firing, while firing became phasic as soon as a threshold was exceeded. The spatial response characteristics of thermal neurons was

studied by Goris and Terashima (1973). Two types of thermal neurons were found (cold and hot), which responded primarily as contrast detectors. This behavior is particularly important for discriminating prey that is either hotter or colder than the background, even while the snake is moving.

Development of the optic tectum is directly related to the diurnal habits of the snake. Tectal lobes are highly developed for diurnal activity, and not so well developed for nocturnal activity. Mapping of the optic tectum and the existence of AND/OR bimodal neurons was demonstrated by Hartline et al. (1978). The AND neurons responded to rapidly moving stimuli in the visual and infrared, while the OR neurons responded to static or moving stimuli in either or both channels. A center-surround organization was also seen, which was theorized to be hot and cold thermal neurons feeding into the optic tectum (Hartline et al., 1978). Thermal neuron properties were measured by Terashima and Liang (1991), who also mapped dendritic structure using HRP labeling. Their investigations found that all paths from the thermal pit receptor terminated in the LTTD.

The existence of six types of bimodal neurons was revealed through the experimental studies of Newman and Hartline (1981). These six types included the AND/OR neurons seen in the previous study (Hartline et al., 1978), along with neurons that exhibited cross-modality interactions. These cross-modality neurons were of the enhanced and inhibitory types. The infrared-enhanced visual neurons responsed weakly to a visual stimulus presented alone, but responded quite strongly to both visual and infrared stimuli presented together, with the corresponding behavior for the visual-enhanced infrared neurons. The infrared-inhibited visual neurons responded very strongly to a visual stimulus presented alone, but lost all response if either an infrared stimulus was presented alone or both stimuli were presented. The cross-modality neurons of most use for saturated sensors are the inhibited type, since target signatures which may be present if only weakly in one sensor modality may be overwhelmed by the saturated sensor.

Fuzzy self-organizing feature maps (Huntsberger & Ajjimarangsee, 1990), hereafter referred to as FSOFM, are used for the initial and final stages of the system shown in Figure 1. This type of network, shown in Figure 2, consists of three layers. The input layer feeds forward into a distance layer which determines the distance between the input vector and the current weights using a predefined metric. The distance layer then feeds forward into a membership layer which calculates the membership value of the input vector to the set of all output vectors. This is done using the following form derived from the fuzzy c-means algorithm (Bezdek, 1981) and

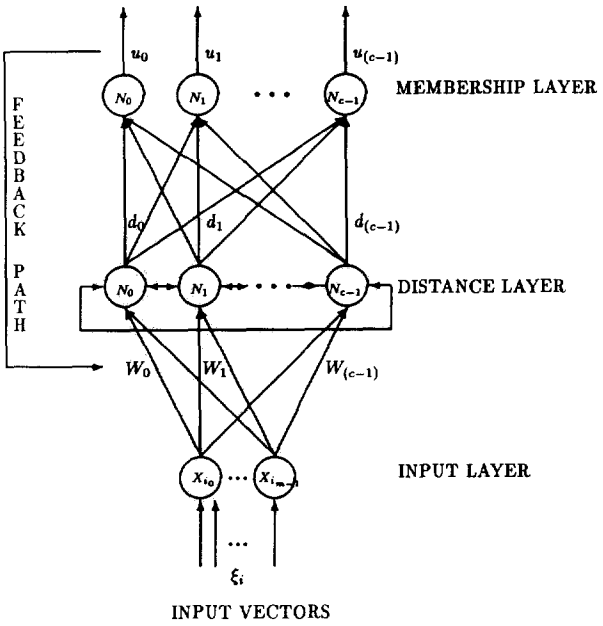

FIGURE 2. Self-organizing feature map with feedback.

used in a previously developed computer vision system (Huntsberger et al., 1985):

$$u_{ji} = \frac{1}{\left[\sum_{l=0}^{c-1} \left(\frac{d_{ji}}{d_{li}}\right)^{\frac{1}{(m-1)}}\right]} \tag{1}$$

where d_{ji} is the distance between the input vector $\boldsymbol{\xi}_i$ and weight vector $\mathbf{W}_j$, c is the number of distance neurons, and m is a weighting exponent which is set within the range 2.0 to ∞. These membership values u_{ji} are then fed back into the network and participate in the weight update rule as:

$$\mathbf{W}_j = \mathbf{W}_j + u_{ji} * dw_j \tag{2}$$

where $dw_j = (\boldsymbol{\xi}_i - \mathbf{W}_j)$. Details and performance of sequential and parallel versions of the algorithm can be found in the original paper (Huntsberger & Ajjimarangsee, 1990). Experimental studies indicated that the network reproduced the segmentation results from an earlier standard implementation of the fuzzy c-means algorithm (Huntsberger et al., 1985; 1986) with close to two orders of magnitude increase in performance. The network was used in a previous study for unsupervised identification of target signatures in multiband imagery (Huntsberger, 1992c).

The optic tectum portion of the system is found in the intermediate stage, where the six types of filters found experimentally by Newman and Hartline (1981) were used to train multiple FSOFM. Since Newman and Hartline only measured response in the absence or presence of a stimulus, the clamped

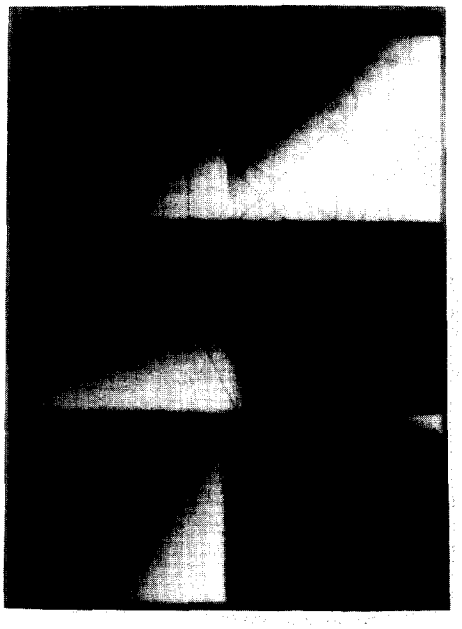

FIGURE 3. Six rattlesnake fusion neuron responses (see text for description).

network was used to generalize to the full range of possible inputs. More will be said about this choice in the experimental section.

After the learning phase of the model has been completed, a lookup table is used for speed of processing. These lookup tables are shown in Figure 3, where for each type of filter, low inputs in both channels correspond to the upper left hand corner of the filter, and high inputs in both channels correspond to the lower right hand corner of the filter. Output from this stage is passed to another FSOFM in the third stage as fractal signature texture vectors (Huntsberger et al., 1987) which are derived from spatial neighborhood sampling of the images. The size of the neighborhood is dependent on the desired resolution for target discrimination (taken to be 9×9 in our experimental studies). This use of fractal signature texture vectors was previously shown to give good segmentation results for multi-channel images (Vafaimagden, 1989).

3. THERMAL NEURON MODELING

Mobile targets present the problem of time-varying stimulus. The spatiotopic mapping of both the thermal and visual portions of the optic tectum are relatively well matched in the "strike" zone in front of the rattlesnake (Hartline et al., 1978). Most of the non-linear portions of the maps are in the posterior region and in the peripheral portions of the thermal sensor. This indicates that processing in the LTTD prior to the optic tectum may provide some clues to the temporal properties of the rattlesnake response. The temporal properties of thermal neurons in the LTTD were recently measured by Terashima and Liang (1991). These parameters can be combined with the earlier work of Goris and Terashima (1973) to give the operating characteristics of thermal neurons under different levels of stimulus.

In our model, spatiotopic neuronal behavior is approximated using a 2-D grid of point soma model neurons. Dendritic structure is captured in extracellular linkage of the outputs. The governing equations for the point soma model (MacGregor, 1987) are:

$$\frac{dE}{dt} = \frac{-E + [P + G_K * (E_K - E)]}{\tau_{LAT}}, \tag{3}$$

$$\frac{dT}{dt} = \frac{-(T - T_o) + c * E}{\tau_T}, \tag{4}$$

$$S = \begin{cases} 0 & \text{if } E < T \\ 1 & \text{otherwise} \end{cases}, \tag{5}$$

$$\frac{dG_K}{dt} = \frac{-G_K + B * S}{\tau_K}, \tag{6}$$

where E is the transmembrane potential, P is the applied potential, G_K is the neuronal potassium conductance, $E_K = 8.35\,\text{mV}$ is the potassium equilibrium potential, $\tau_{LAT} = 0.8\,\text{ms}$ is the latency time, T is the adaptive firing threshold, $c = 0.75$ is the threshold rise, $T_o = -55.13\,\text{mV}$ is the resting threshold of the cell, $\tau_T = 1.76\,\text{ms}$ is the time constant rise of the threshold, S is the spike generation parameter, $B = 20\,\text{m mho/cm}^2$ is the post-firing potassim conductance rise, and $\tau_K = 3.81\,\text{ms}$ is the potassium decay time constant. The values for these parameters used in our simulation are taken from the work of Terashima and Liang (1991) and from MacGregor (1987).

The measured latency time for thermal neurons varied from 50 ms to no response in the studies of Goris and Terashima (1973), depending on the strength of the input stimulus. This means that a stimulus with a strong contrast to the background will cause rapid firing of neurons, with a highly non-linear decrease in firing frequency as the contrast decreases. This effect can be included in the model using an exponential weighting of τ_{LAT} of the form

$$\tau'_{LAT} = \tau_{LAT} e^{((1.0-P)/2.0)}, \tag{7}$$

for normalized input stimuli between 0.5 and 1.0, and

$$\tau'_{LAT} = \tau_{LAT} e^{((1.0-P)/1.05)}, \tag{8}$$

for normalized stimuli less than 0.5, with the latency

FIGURE 4. Accommodation of thermal neuron array to a static stimulus.

time measured with direct electrical stimulation is given by τ_{LAT}.

Rapid accommodation to a static input scene occurs in the rattlesnake, and the AND fusion neurons usually only respond to rapidly moving stimuli (Hartline et al., 1978). This response is of use to the rattlesnake to filter out background clutter, since prey will usually be moving. Figure 4 shows a temporal study of the thermal neuron array that is excited by a static stimulus. The time step between frames is 50 ms. The gray scale in this series represents the transmembrane potential, with dark for negative values and light for positive values. We have used pacemaker neurons in order to see the areas of uniform response more clearly. By the fourth frame the thermal neuron array has accommodated to the pattern, and feature discrimination is no longer possible. Without any motion in the field of view, object discrimination is not possible with an array of thermal neurons.

The rattlesnake offsets this effect by slightly moving its head when it is trying to distinguish prey from background. This causes a slight motion of the field of view across the thermal neuron field. Motion tracking occurs through the highly non-linear latency of the thermal neuron response. Prey with sufficient contrast will cause rapid firing of neurons that lie within the boundaries and leading edge that are projected onto the thermal neuron array. This behavior is accompanied by a trail of low firing rate neurons in the areas of the array that are disoccluded due to apparent motion of the prey. This effect also carries over to objects that are undergoing actual motion across the thermal neuron array. Analogous behavior was seen in motion analysis systems that replicated primate visual responses (Hutchinson et al., 1988; Grossberg & Rudd, 1989; Seibert & Waxman, 1989; Marshall, 1990). This effect can be seen in Figure 5, where a rectangular object is moving from left to right across a field of neurons whose states have been randomly initialized. The random initialization is based on aperiodic firing of thermal neurons in the absence of a specific stimulus, as reported by Bullock and Diecke (1956), and Goris and Terashima (1973). The object causes the neurons within its projected boundaries to fire, with a diffuse trail of neuronal activity behind it. The next section reports some experimental studies on real-world data of the static and motion systems seen in this and the previous section.

4. EXPERIMENTAL STUDIES

We performed two series of experiments using multichannel infrared data from a test range at Fort Drum, New York. The three infrared channels are polarization, reflectance and thermal. The images were obtained from a helicopter flyover of the site. The multichannel inputs were obtained at four times during the day (0000, 0600, 1200, and 1800 hours). All of the targets are fixed, and the area of each is about 0.003% of the field of view. The images are spatially registered for each time of day, but not across the times. We decided to use the polarization

FIGURE 5. Response of thermal neuron array to dynamic stimuli.

FIGURE 6. Three channel infrared scene (0000 hours).

and thermal channels as the bimodal inputs, since the reflectance channel contained very little information due to its limited dynamic range.

The first study was designed to test the sensitivity of the fusion system to diurnal variations in the input. All runs for the first study were done using the parallel version of the FSOFM algorithm (Huntsberger & Ajjimarangsee, 1990) on 32 nodes of the Intel Paragon at the University of South Carolina. The time for a two channel analysis took 3.6 s. Based on previous scaling studies performed on the 1024 node nCUBE at the University of South Carolina (Hunbsberger & Ajjimarangsee, 1990), this time would be reduced to 226 ms on a 512 node Paragon. This puts the algorithm within real-time performance range on a general purpose supercomputer.

A small subsample of the reflectance, polarization, and thermal channel images are shown in Figures 6–9 for each time of day. The targets are visible in the thermal channels at 0600, 1200 and 0000 hours due to their differential heating rate compared to the background, but heating during the day has almost saturated the thermal channel at 1800 hours. There is virtually no possibility of recovering target locations during this time period without using another one of the channels. This situation is similar to the one that a rattlesnake would find itself in during the same time of day. The pit organ would be close to saturated, and visible sensing would be used with the cross-modality neurons to disambiguate the scene.

The classification results are shown in Figures 10–13 for the four times of day. In these figures, the targets that are found are enclosed by boxes, and the arrows point to false-alarms if boxed or misses if not. We have used the thermal channel images for the labeled output, except in the 1800 hour sample where we have used the polarization channel for clarity. We compared the P_D and FAR from each of the two channels (polarization and thermal) to the fused channel analysis. Single channel results are derived using just the first and third stages of the system, without the fusion filters. The P_D calculated using ground truth information for the study are shown in Table 1, and the FAR are shown in Table 2. All P_D

FIGURE 7. Three channel infrared scene (0600 hours).

FIGURE 8. Three channel infrared scene (1200 hours).

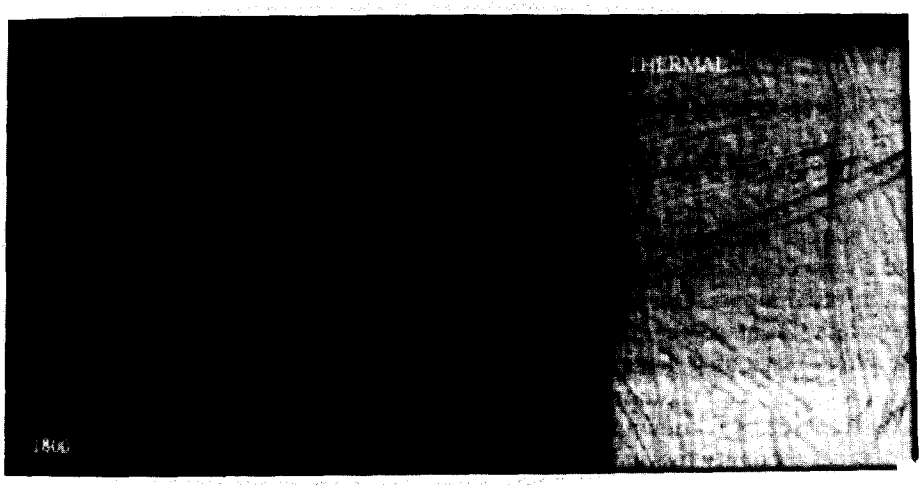

FIGURE 9. Three channel infrared scene (1800 hours).

FIGURE 10. Target detection results from fused polarization and thermal infrared channels (0000 hours).

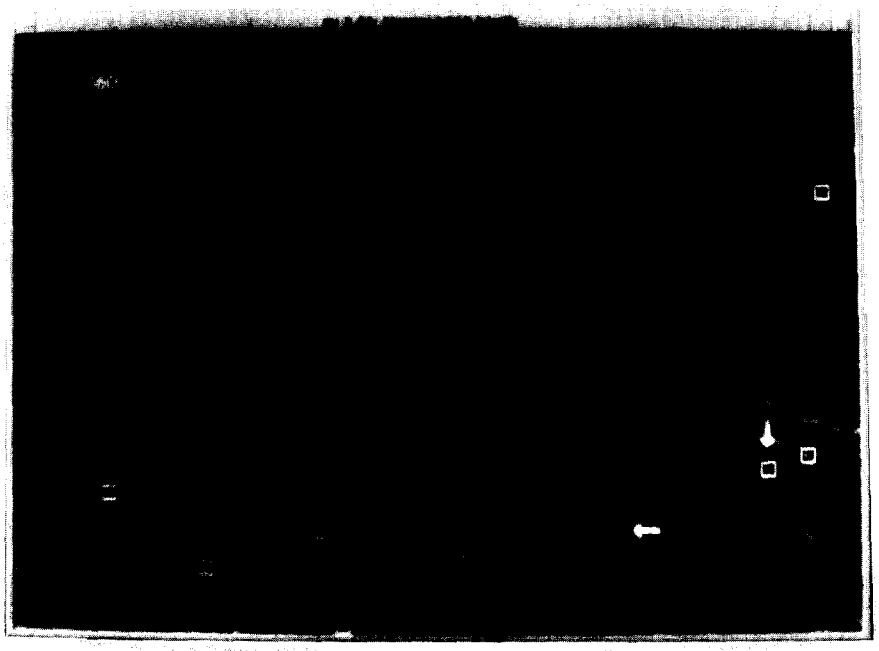

FIGURE 11. Target detection results from fused polarization and thermal infrared channels (0600 hours).

are given in percent and all FAR are given in $10^6\times$ false-alarms/km^2.

The performance is quite remarkable, considering that the fusion filters were developed using data from visible and thermal infrared inputs. Our other studies on visible and range data (Huntsberger, 1990, 1992a) indicate that these filters serve as a general-purpose set for any cross-modality fusion process. If the sensor response characteristics are known for certain target types, a set of cross-modality fusion filters with non-linear behavior can be custom designed and substituted for the middle stage of the fusion system. For example, in the case of targets which respond weakly to one sensor and strongly to another, the FSOFM generalization for the INFRINH and VISINH filters shown in Figure 3 can be stretched toward the origin to favor this combination. In addition, adaptive filters can be designed that respond to changes in environmental conditions.

The relatively high P_D seen in Table 1 for the thermal channel taken alone is accompanied by FAR that are close to two orders of magnitude worse than

FIGURE 12. Target detection results from fused polarization and thermal infrared channels (1200 hours).

FIGURE 13. Target detection results from fused polarization and thermal infrared channels (1800 hours).

the fused channel results. Particularly of note is the drop to no false-alarms in the 1200 hour and 1800 hour sampled images. The FAR are as good or better than the results of Yu et al. (1993), which is one of the best statistical methods for ATR in terms of P_D and FAR. Their studies were done on thematic mapper (TM) multiband images taken at a much higher altitude than our study region. Target spatial and spectral properties for one time of day were used to build adaptive templates for identification. The performance of our system was achieved without the need to fine-tune the spectral and spatial profiles of the targets for each time of day, since the combination of fusion filters and processing in the initial stage acted as a type of normalization.

The second experimental study investigated the use of the thermal neuron array for mobile target discrimination. We used the 1800 hour thermal channel image to simulate a mobile target as would be sensed by a moving platform, such as a missile or helicopter. Four frames from the sequence are shown in Figure 14, where the mobile target is translating diagonally from left to right and the platform is translating in the same direction. This would be analogous to the case of a moving platform matching a mobile target's direction and then tracking. The rates of translation are 37.5 pixels/frame for the target and 66.6 pixels/frame for the platform.

The response of the array to the four thermal images is shown in Figure 15, where the same neuronal parameters are used as in the discussion from the previous section. At first, the thermal neuron array is saturated, but by the second frame it has started to accommodate to the moving background clutter. Although there is a lot of ground clutter that is moving at a different rate relative to the target, target discrimination is fairly good. This property of the thermal neuron array occurs for any low contrast background, whether stationary or moving.

There are some limitations to the effectiveness of the thermal neuron array in its current form for target discrimination. Since the model only has nearest neighbor simulation of the extracellular linkage, very fast motion of multiple targets relative to the potassium decay time constant τ_K of 3.81 ms would not give the array time to recover between excitation events. This can be addressed in multiple

TABLE 1
P_D(%) for Four Times of Day and Various Channel Configurations

Sensor	Time of Day			
	0000	0600	1200	1800
Polar	45.7	50.5	60.8	71.5
Therm	85.4	89.1	80.5	24.3
Fused	93.8	100.0	90.9	92.9

TABLE 2
FAR(10^6 × false-alarms/km^2) for Four Times of Day and Various Channel Configurations

Sensor	Time of Day			
	0000	0600	1200	1800
Polar	85.6	94.3	102.8	57.5
Therm	100.4	109.5	87.7	975.6
Fused	7.4	8.8	0.0	0.0

FIGURE 14. Four frames from thermal infrared motion sequence.

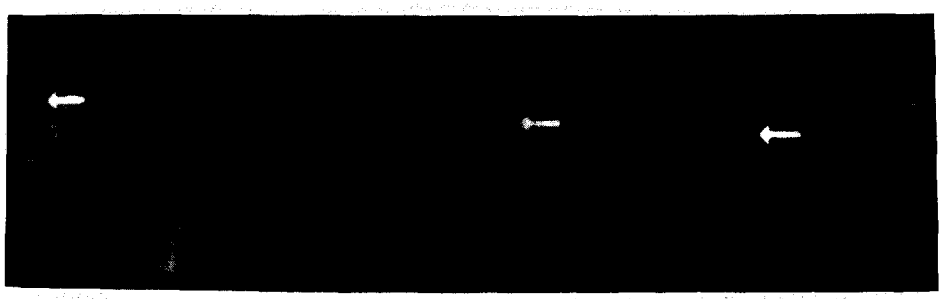

FIGURE 15. Thermal neuron array response to thermal infrared motion sequence.

ways: by decreasing τ_K, increasing the time constant rise of the threshold τ_T, or a combination of both. The trade-off here is the possibility of an increased number of false firings, since the threshold takes a longer time to clamp.

The clutter/contrast response of the thermal neuron array is controlled by our modified form of the latency time τ'_{LAT}, parameters of which were determined using the experimental rattlesnake data. The exponential form enhances the leading and trailing edges of the array response to a moving target, but can have an adverse effect for multiple targets. Targets following in the wake would tend to be obscured by the slow firing rate of the neurons. This problem can be addressed by using a functional form of τ'_{LAT} that peaks at a slower rate. In addition, the longer range extracellular interactions from a more complete dendritic model would minimize these effects. Further studies will be needed to determine the optimal parameters needed for any given scenario.

The position of the target within the 2-D array differs from the actual target position in the original images due to the relatively large velocity components of the platform and mobile target. Despite the obscuration of the target when its signature merges with the ground clutter, as is evidenced in the third frame in the sequence in Figure 14, the thermal neuron array still acts as a discriminator. Based on the results of the first experimental study, the use of other channels in a fused analysis would serve to further enhance the target signature. In addition, the texture of the region surrounding the moving target differs from that of the moving background clutter, and would be able to be distinguished using the fractal signature approach found in the full fusion system.

5. CONCLUSIONS

We have presented a multi-stage neural network model for automatic target recognition. The neural network model is based on neurons found in the optic tectum of the rattlesnake which directly integrate visible and thermal infrared sensory inputs. The six filters from the middle stage of the network have shown themselves to be suitable for integration of multichannel infrared sensory inputs. Based on our experimental studies, near real-time performance (226 ms/frame) of the system can be achieved on a 512 node Intel Paragon. The cross-modality nature of the filters tended to minimize the effect that diurnal variations in the inputs had on target classification. At 1800 hours in the Ft. Drum sequence, the thermal infrared channel was almost completely saturated due to heating of the scene throughout the day. This channel, when fused with the polarization channel, returned a P_D of 92.9% with no false-alarms. In all cases the P_D are greater than 90% for our test sequence, and the FAR are sometimes two orders of magnitude better than discrimination based on single channel inputs.

The FSOFM used in the first and final stages of the fusion system shown in Figure 1 is not limited to two input sensors. Treatment of multispectral data is straightforward, as demonstrated in previous studies where multiband images were segmented without

using the fusion tables (Huntsberger & Ajjimarangsee, 1990). The only limitation in the current system is the need for registration of the sensor output images. For a full fusion analysis, the lookup tables would have to be modified, or used in a cascade mode to handle the multiple inputs.

We also presented some preliminary results using a 2-D array of rattlesnake thermal neurons for mobile target discrimination. Our initial studies indicate that use of the array with a single sensor input yields low level motion detection even in the presence of ground clutter. We are presently investigating the use of the fusion system for multichannel analysis of the motion sequences. The first stage of the fusion system will be replaced with the thermal neuron array in order to compare its contrast enhancement to that of the FSOFM. Results from the first experimental study indicate that much better target discrimination is possible using this approach. We are also examining tracking methods for the 2-D array response patterns.

REFERENCES

Ajjimarangsee, P., & Huntsberger, T. L. (1988). Neural network model for the fusion of visible and thermal infrared sensor output. In P. Schenker (Ed.), *Proceedings of SPIE Symposium on Sensor Fusion I* (Vol. 1003, pp. 153–160). Cambridge, MA.

Bai, B., & Farhat, N. H. (1992). Learning networks for extrapolation and radar target identification. *Neural Networks*, **5**, 507–529.

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective functon algorithms*. New York: Plenum Press.

Bullock, T. H., & Diecke, F. P. J. (1956). Properties of an infra-red receptor. *Journal of Physiology*, **134**, 47–87.

Chatterjee, P. S., & Huntsberger, T. L. (1988). Comparison of techniques for sensor fusion under uncertain conditions. In P. Schenker (Ed.), *Proceedings SPIE Symposium on Sensor Fusion I* (Vol. 1003, pp. 194–199). Cambridge, MA.

Chu, C. C., & Aggarwal, J. K. (1992). Image interpretation using multiple sensing modalities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14**, 841–847.

Farhat, N. H., & Bai, B. (1989). Echo inversion and target shape estimation by neuromorphic processing. *Neural Networks*, **2**, 117–125.

Goris, R. C., & Terashima, S. (1973). Central response to infra-red stimulation of the pit receptors in a Crotaline snake, *Trimeresurus flavoviridis*. *Journal of Experimental Biology*, **58**, 59–76.

Gorman, R. P., & Sejnowski, T. J. (1988). Learned classification of sonar targets using a massively parallel network. *IEEE Transactions on Acoustic, Speech, and Signal Processing*, **36**, 1135–1140.

Grossberg, S., Guenther, F., Bullock, D., & Greve, D. (1993). Neural representations for sensory-motor control, II: Learning a head-centered visuomotor representation of 3-D target position. *Neural Networks*, **6**, 43–67.

Grossberg, S., & Rudd, M. E. (1989). A neural architecture for visual motion perception: Group and element apparent motion. *Neural Networks*, **2**, 421–450.

Hartline, P. H., Kass, L., & Loop, M. S. (1978). Merging of modalities in the optic tectum: Infrared and visual integration in rattlesnakes. *Science*, **199**, 1225–1229.

Hemminger, T. L., & Pao, Y.-H. (1994). Detection and classification of underwater acoustic transients using neural networks. *IEEE Transactions on Neural Networks*, **5**, 712–718.

Huntsberger, T. L. (1990). Comparison of techniques for disparate sensor fusion. In *Proceedings of SPIE Symposium on Sensor Fusion III: 3-D Perception and Recognition* (Vol. 1383, pp. 589–594). Boston, MA.

Huntsberger, T. L. (1992a). Data fusion: A neural networks implementation. In M. A. Abidi & R. C. Gonzalez (Eds.), *Data fusion in robotics and machine intelligence*. Orlando, FL: Academic Press.

Huntsberger, T. L. (1992b). Sensor fusion in a dynamic environment. In P. Schenker (Ed.), *Proceedings of SPIE Symposium on Sensor Fusion V*. (Vol. 1828, pp. 175–182). Boston, MA.

Huntsberger, T. L. (1992c). Application of parallel self-organizing neural networks to automatic target cueing of multiple band imagery. Final Report, ARO Contract DAAL03-91-C-0034, January, 1992.

Huntsberger, T. L., & Jayaramamurthy, S. N. (1987a). Determination of the optic flow field using the spatiotemporal deformation of region properties. *Pattern Recognition Letters*, **6**, 169–177.

Huntsberger, T. L., & Jayaramamurthy, S. N. (1987b). A framework for multi-sensor fusion in the presence of uncertainty. In *Proceedings of Workshop on Spatial Reasoning and Multi-sensor Fusion* (pp. 345–350). St. Charles, IL.

Huntsberger, T. L., & Jayaramamurthy, S. N. (1988). Determination of the optic flow field in the presence of occlusion. *Pattern Recognition Letters*, **8**, 325–333.

Huntsberger, T. L., & Ajjimarangsee, P. (1990). Parallel self-organizing feature maps for unsupervised pattern recognition. *International Journal of General Systems*, **16**, 357–372. Also reprinted in *Fuzzy models for pattern recognition*. J. C. Bezdek & S. K. Pal (Eds.) (pp. 483–495). Piscataway, NJ: IEEE Press.

Huntsberger, T. L., Jacobs, C. L., & Cannon, R. L. (1985). Iterative fuzzy image segmentation. *Pattern Recognition*, **18**, 131–138.

Huntsberger, T. L., Rangarajan, C., & Jayaramamurthy, S. N. (1986). Representation of uncertainty in computer vision using fuzzy sets. *IEEE Transactions on Computers*, **35**, 145–156. Also reprinted in *Fuzzy models for pattern recognition* J. C. Bezdek & S. K. Pal (Eds). (pp. 397–408). Piscataway, NJ: IEEE Press. 1992.

Huntsberger, B. A., Jayaramamurthy, S. N., & Huntsberger, T. L. (1987). Morphological classification of color textures. In *Proceedings of IEEE 30th Midwest Symposium on Circuits and Systems* (pp. 363–366). Syracuse, New York.

Hutchinson, J., Koch, C., Luo, J., & Mead, C. (1988). Computing motion using analog and binary resistive networks. *IEEE Computer*, **21**, 52–64.

Ishizuka, M. (1981). Extension of Dempster and Shafer's theory to fuzzy set for constructing expert systems. Summary of Papers on General Fuzzy Problems 7, Working group on fuzzy systems, Tokyo, Japan.

Krishnapuram, R., & Lee, J. (1992). Fuzzy-set-based hierarchical networks for information fusion in computer vision. *Neural Networks*, **5**, 335–350.

Landowski, J. G., & Fong, D. Y.-S. (1990). A self-organizing neural network and its potential application in an autonomous target recognition system. In *Proceedings of the Lockheed Intelligent Systems Symposium*, Palo Alto, CA.

MacGregor, R. J. (1987). *Neural and brain modeling*. San Diego, CA: Academic Press.

Marshall, J. A. (1990). Self-organizing neural networks for perception of visual motion. *Neural Networks*, **3**, 45–74.

Nandhakumar, N., & Aggarwal, J. K. (1988). Integrated analysis of thermal and visual images for scene interpretation. *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, **10**, 469–481.

Newman, E. A., & Hartline, P. H. (1981). Integration of visual and infrared information in bimodal neurons of the rattlesnake tectum. *Science*, **213**, 789–791.

Paulin, M. G., Nelson, M. E., & Bower, J. M. (1989). Neural control of sensory acquisition: The vestibulo-ocular reflex. In D. S. Touretzky (Ed.), *Advances in neural information processing systems* (Vol. I, pp. 410–418) CA: Morgan-Kaufmann.

Perlovsky, L. I. (1987). Multiple sensor fusion and neural networks. In *DARPA neural network study*. Lexington, MA: MIT/Lincoln Laboratory.

Perlovsky, L. I., & McManus, M. M. (1991). Maximum likelihood neural networks for sensor fusion and adaptive classification. *Neural Networks*, **4**, 89–102.

Roth, M. W. (1989). Neural networks for extraction of weak targets in high clutter environments. *IEEE Transactions on Systems, Man, and Cybernetics*, **19**, 1210–1217.

Roth, M. W. (1990). Survey of neural network technology for automatic target recognition. *IEEE Transactions on Neural Networks*, **1**, 28–43.

Schiff, H., Boarino, P. C., Corso, D. D., & Filippi, E. (1994). A hardware implementation of a biological neural system for target localization. *IEEE Transactions on Neural Networks*, **5**, 354–362.

Seibert, M., & Waxman, A. M. (1989). Spreading activation layers, visual saccades, and invariant representations for neural pattern recognition systems. *Neural Networks*, **2**, 9–20.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press.

Simmons, J. A. (1990). Acoustic-imaging computations by echolocating bats: Unification of diversely-represented stimulus features into whole images. In D. S. Touretzky (Ed.), *Advances in neural information processing systems* (Vol. II, pp. 2–9). CA: Morgan-Kaufmann.

Spence, C. D., & Pearson, J. C. (1990). The computation of sound source elevation in the barn owl. In D. S. Touretzky (Ed.), *Advances in neural information processing systems* (Vol. II, pp. 10–17). CA: Morgan-Kaufmann.

Spence, C. D., Pearson, J. C., Gelfand, J. J., Peterson, R. M., & Sullivan, W. (1989). Neuronal maps for sensory-motor control in the barn owl. In D. S. Touretzky (Ed.), *Advances in neural information processing systems* (Vol. I, p. 366–374). CA: Morgan-Kaufmann.

Terashima, S., & Liang, Y. (1991). Temperature neurons in the Crotaline trigeminal ganglia. *Journal of Neurophysiology*, **66**, 623–634.

Vafaimagden, S. (1989). Texture based image segmentation. Master's thesis, University of South Carolina, Department of Computer Science.

Yu, X., Reed, I. S., & Stocker, A. D. (1993). Comparative performance analysis of adaptive multi-spectral detectors. *IEEE Transactions on Signal Processing*, **41**(8),2639–2656.